

Human Motion Diffusion as a Generative Prior

Year: 2023 | Citations: 310 | Authors: Yonatan Shafir, Guy Tevet, Roy Kapon, Amit H. Bermano

Abstract

Recent work has demonstrated the significant potential of denoising diffusion models for generating human motion, including text-to-motion capabilities. However, these methods are restricted by the paucity of annotated motion data, a focus on single-person motions, and a lack of detailed control. In this paper, we introduce three forms of composition based on diffusion priors: sequential, parallel, and model composition. Using sequential composition, we tackle the challenge of long sequence generation. We introduce DoubleTake, an inference-time method with which we generate long animations consisting of sequences of prompted intervals and their transitions, using a prior trained only for short clips. Using parallel composition, we show promising steps toward two-person generation. Beginning with two fixed priors as well as a few two-person training examples, we learn a slim communication block, ComMDM, to coordinate interaction between the two resulting motions. Lastly, using model composition, we first train individual priors to complete motions that realize a prescribed motion for a given joint. We then introduce DiffusionBlending, an interpolation mechanism to effectively blend several such models to enable flexible and efficient fine-grained joint and trajectory-level control and editing. We evaluate the composition methods using an off-the-shelf motion diffusion model, and further compare the results to dedicated models trained for these specific tasks.